WATER RESOURCES RESEARCH, VOL. 48, W07525, doi:10.1029/2011WR011291, 2012

Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations

Rasmus Houborg,[1,2] Matthew Rodell,[3] Bailing Li,[2,3] Rolf Reichle,[3] and Benjamin F. Zaitchik[4]

Received 17 August 2011; revised 1 June 2012; accepted 12 June 2012; published 28 July 2012.

[1] The Gravity Recovery and Climate Experiment (GRACE) twin satellites observe time variations in Earth's gravity field which yield valuable information about changes in terrestrial water storage (TWS). GRACE is characterized by low spatial (>150,000 km^2) and temporal (>10 days) resolution but has the unique ability to sense water stored at all levels (including groundwater) systematically and continuously. The GRACE Data Assimilation System (DAS), based on the Catchment Land Surface Model (CLSM), enhances the value of the GRACE water storage data by enabling spatial and temporal downscaling and vertical decomposition into moisture components (i.e., groundwater, soil moisture, and snow), which individually are more useful for scientific applications. In this study, GRACE DAS was applied to North America, and GRACE-based drought indicators were developed as part of a larger effort to investigate the possibility of more comprehensive and objective identification of drought conditions by integrating spatially, temporally, and vertically disaggregated GRACE data into the U.S. and North American Drought Monitors. Previously, the drought monitors lacked objective information on deep soil moisture and groundwater conditions, which are useful indicators of drought. Extensive data sets of groundwater storage from U.S. Geological Survey monitoring wells and soil moisture from the Soil Climate Analysis Network were used to assess improvements in the hydrological modeling skill resulting from the assimilation of GRACE TWS data. The results point toward modest, but statistically significant, improvements in the hydrological modeling skill across major parts of the United States, highlighting the potential value of a GRACE-assimilated water storage field for improving drought detection.

Citation: Houborg, R., M. Rodell, B. Li, R. Reichle, and B. F. Zaitchik (2012), Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, *Water Resour. Res.*, *48*, W07525, doi:10.1029/2011WR011291.

1. Introduction

[2] Accurate characterization of the timing, duration, and severity of drought events is extremely important because of droughts' potentially devastating impacts on society. Droughts affect more people than any other natural hazard [*Wilhite*, 2000] and the National Drought Mitigation Center (NDMC) has estimated the average annual economic loss in the United States due to drought at $6–8 billion, more than any other type of disaster [*Western Governors' Association*, 2004].

[3] The U.S. and North American Drought Monitors [*Svoboda et al.*, 2002; *Lawrimore et al.*, 2002] have been successful in defining, monitoring, and predicting drought and have proven to be valuable tools available to water resource managers and decision makers for assessing and mitigating drought impacts and for reducing the vulnerability of society to drought. However, the drought monitors (DM) rely heavily on precipitation indices and subjective information, and they do not currently incorporate systematic observations of subsurface water storage because of the scarcity of reliable and objective information. Groundwater levels in near-surface unconfined or semiconfined aquifers are particularly well suited to drought monitoring because groundwater storage reflects meteorological conditions occurring over timescales of days to years, whereas near-surface water stores respond more quickly to rainfall (or lack thereof). In other words, on the timescales of droughts, which generally develop over a period of weeks to months and can last for a decade or more, groundwater levels respond (absent other influences) somewhat proportionally to the degree and persistence of anomalous hydroclimatic conditions [*Mishra and Singh*, 2010]. Further, groundwater is a valuable resource that is often strained when other sources of water are scarce. However, continuous observation of large-scale groundwater storage in space

[1]Climate Risk Management Unit, Institute for Environment and Sustainability, Joint Research Centre, European Commission, Ispra, Italy.
[2]Earth System Science Interdisciplinary Center, College Park, Maryland, USA.
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[4]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA.

Corresponding author: R. Houborg, Climate Risk Management Unit, Institute for Environment and Sustainability, Joint Research Centre, European Commission, Via Enrico Fermi 2749, I-21027 Ispra, Italy. (rasmus.houborg@jrc.ec.europa.eu)

0043-1397/12/2011WR011291

and time is impaired by inadequate well networks and limited capacity to quantify groundwater storage from space.

[4] Satellite observations of the Earth's time-variable gravity field from the Gravity Recovery and Climate Experiment (GRACE) mission [*Tapley et al.*, 2004] enable reliable detection of variations in total terrestrial water storage (TWS; i.e., the sum of groundwater, soil moisture, snow, surface water, ice, and biomass), thereby addressing the observational gap of monitoring regional to continental scale water storage changes on a systematic basis. Since its launch in March 2002, GRACE has provided unprecedented observations of water storage dynamics at the basin to continental scale [*Wahr et al.*, 2006], which have improved quantification and understanding of hydrologic states and fluxes at regional to global scales; GRACE observations have been used for inferring terrestrial freshwater discharge [*Syed et al.*, 2008], evapotranspiration [*Boronina and Ramillien*, 2008; *Rodell et al.*, 2004; *Ramillien et al.*, 2006; *Sheffield et al.*, 2009], the mass balance of ice sheets [*Chen et al.*, 2009a; *Velicogna*, 2009] and glaciers [*Luthcke et al.*, 2008], and the water balance of lakes [*Swenson and Wahr*, 2009]. Groundwater [*Rodell et al.*, 2007, 2009; *Strassberg et al.*, 2009; *Swenson et al.*, 2008; *Yeh et al.*, 2006] and snow [*Niu et al.*, 2007] have been isolated from GRACE TWS using auxiliary information. GRACE-based storage changes are in good agreement with those obtained from land surface model (LSM) simulations [*Güntner*, 2008; *Syed et al.*, 2008] and in situ observations [*Rodell et al.*, 2007; *Swenson et al.*, 2006; *Syed et al.*, 2005; *Yeh et al.*, 2006], and the utility of GRACE for characterizing extreme drought has been demonstrated in a number of recent studies [*Yirdaw et al.*, 2008; *Chen et al.*, 2009b; *Leblanc et al.*, 2009]. Thus the potential to use GRACE observations to fill the current need for subsurface water information in the drought mapping process is evident.

[5] While GRACE has supported many advances in water cycle science, the monthly production frequency and coarse spatial resolution ($\sim$150,000 km^2 [*Rowlands et al.*, 2005; *Swenson et al.*, 2006]) limit the utility of GRACE observations for a majority of applications that require near-real-time input of much finer resolution earth observation data. In order to realize the full potential of GRACE for hydrological applications the basin-scale, column-integrated, monthly TWS anomalies from GRACE must be effectively downscaled in space and time, vertically stratified into moisture component anomalies (e.g., soil moisture, groundwater, snow), and extrapolated to the present, thereby meeting the specificity, timeliness, and high spatial resolution requirements of most applications.

[6] Data assimilation, which synthesizes the advantages of observations and numerical land surface models, can be used to disaggregate GRACE observations temporally, horizontally, and vertically. This was demonstrated by *Zaitchik et al.* [2008], who assimilated GRACE TWS anomalies into the Catchment LSM (CLSM) using a novel implementation of an Ensemble Kalman Smoother. This GRACE Data Assimilation Scheme (GRACE DAS) was shown to improve model skill in the simulation of hydrological states and fluxes at sub-GRACE resolution in the Mississippi basin [*Zaitchik et al.*, 2008]. Data assimilation serves to reduce uncertainties in LSM simulation resulting from the input data used to force the LSMs, simplifications in model parameterization and limitations in the described physical realism of the model, by using observation data sets for constraining LSM simulations of terrestrial hydrology. GRACE-based hydrological data clearly have the potential for improving global LSMs [*Güntner*, 2008; *Lo et al.*, 2010; *Werth et al.*, 2009]. Other studies have highlighted limitations in the ability of current models to reproduce the amplitudes in TWS change observed by GRACE [*Schmidt et al.*, 2006; *Hasegawa et al.*, 2009], which is partly related to inadequate treatment of groundwater dynamics [*Niu et al.*, 2007; *Rodell et al.*, 2004]. However, so far only very few studies have taken the first steps toward evaluating the benefits of assimilating GRACE TWS data within a LSM [*Zaitchik et al.*, 2008; *Su et al.*, 2010; *Forman et al.*, 2011].

[7] In this study we extend the GRACE DAS of *Zaitchik et al.* [2008] to the North American domain as part of a larger project aimed toward integrating enhanced (i.e., via data assimilation) GRACE TWS data into the U.S. and North American Drought Monitors. Besides the wider range of hydroclimatic conditions, the present study goes beyond *Zaitchik et al.* [2008] by assessing the potential of GRACE DAS for drought monitoring and by evaluating GRACE DAS simulations using soil moisture measurements and groundwater observations beyond the Mississippi River Basin. We expect that drought conditions can be identified more comprehensively and objectively by integrating GRACE-based drought indicators into the short- and long-term "objective blends" (a fusion of precipitation data, various standardized indices such as the Palmer Drought Severity Index, and simple water budget estimates of soil moisture) that constitute U.S. and North American Drought Monitor baselines and currently lack valuable information on deep (root zone and below) soil moisture and groundwater storage changes. Deep soil moisture and groundwater are valuable as drought indicators because they encompass the impact of natural water demand (unlike precipitation indices) and also continue to adjust after surficial drought indicators such as surface soil moisture and vegetation greenness have met their limits of dryness. This paper describes the development of GRACE-based soil moisture and groundwater drought indicators. These are derived from model simulated fields of soil moisture and groundwater constrained by GRACE TWS observations via data assimilation.

[8] Improvements in the hydrologic modeling skill resulting from the GRACE data assimilation is assessed using groundwater storage data from monitoring wells distributed across the United States and soil moisture observations from the SCAN network. The integration of the GRACE-based drought indicators into the operational production of objective DM blends and their effect on drought monitoring by the U.S. and North American Drought Monitors will be the topic of a forthcoming paper.

2. Data and Methods

[9] This section describes the specifics of the adopted GRACE TWS data set (section 2.1), the Catchment Land Surface Model (CLSM; section 2.2), the GRACE Data Assimilation System (section 2.3), and the data used to force and initialize the simulations (section 2.4). A strategy to increase the water storage capacity of the CLSM to better accommodate the GRACE anomalies during times of drought is described in section 2.5, followed by descriptions of the

U.S. and North American Drought Monitors (DM) and the method for generating GRACE-based drought indicator percentiles (section 2.6). The section concludes with a description of the data sets used to evaluate GRACE DAS output (section 2.8).

2.1. GRACE Terrestrial Water Storage

[10] GRACE consists of two satellites in identical orbits about 200 km apart and at 450–500 km altitude, whose velocities respond to changes in gravity. When they approach a positive mass anomaly such as a mountain range, gravitational pull increases and the leading satellite accelerates, increasing the distance between the two, before the second satellite accelerates and catches up. Distance variations between the two satellites are measured precisely using a microwave ranging system and the corresponding time variations in the gravity field are used to determine changes in the Earth's mass distribution at horizontal resolutions no better than $\sim$150,000 km^2, with higher measurement accuracy at larger spatial scales [*Wahr et al.*, 2004]. Monthly to decadal temporal changes in the gravity field are caused mostly by mass redistributions in the atmosphere, ocean and continents. Thus, estimates of changes in total terrestrial water storage (but not the absolute quantity) can be inferred by modeling atmospheric and oceanic circulations and removing their gravitational effects, which are typically much better understood and constrained by observations.

[11] We used GRACE monthly mass grids optimized for land applications covering the period August 2002 to July 2009 and made available at the GRACE Tellus Web site (http://grace.jpl.nasa.gov/data/gracemonthlymassgridsland/). The land product is based on spherical harmonic fields produced by the University of Texas Center for Space Research (CSR; version CSR-RL04). Each terrestrial gravity field solution is composed of a set of spherical harmonic coefficients up to harmonic degree and order 60 that describe the Earth's gravity field with the contributions of solid earth tides, atmosphere, and oceans already removed. Additional postprocessing steps are required to convert the coefficients into maps of TWS anomalies, which include smoothing (300 km half-width Gaussian filter [*Wahr et al.*, 1998]) and removal of correlated errors ("destriping" [*Swenson and Wahr*, 2006]) after subtraction of the long-term temporal mean field (accounting for Earth's static gravity field). A correction for postglacial rebound has also been applied to the data [*Chambers et al.*, 2010]. The spatial averaging (or smoothing) of GRACE data serves to reduce the contribution of noisy short-wavelength components of the gravity field solution. The application of Gaussian and destriping filters are necessary to make geophysical ("real") signals apparent, but at the cost of some degree of signal attenuation [*Klees et al.*, 2006]. In an effort to restore the original signal amplitudes, a grid of multiplicative scaling coefficients has been applied to the monthly land GRACE mass grids. The scaling coefficients were derived independently of the GRACE data by applying the same filtering techniques to modeled TWS data and then computing the signal attenuation at each geographic location (S. Swenson, Restoring signal loss in GRACE terrestrial water storage estimates, manuscript in preparation, 2010, available at http://grace.jpl.nasa.gov/files/swenson.grace.scaling_description_doc_draft.pdf).

2.2. Catchment Land Surface Model

[12] The Catchment Land Surface Model (CLSM) [*Koster et al.*, 2000; *Ducharne et al.*, 2000] is a numerical model that ingests near surface meteorological "forcing" data (e.g., rainfall and downward shortwave radiation) and uses physical equations to determine the evolution of water and energy states (e.g., soil moisture and temperature) and fluxes (e.g., evaporation and sensible heat flux). CLSM abandons the traditional gridded delineation and instead divides the land surface into irregularly shaped catchments, with boundaries defined by topography (Figure 1). This is advantageous when simulating surface hydrological processes as it allows for a more realistic treatment of horizontal heterogeneity in surface properties in contrast to conventional land surface models that assume uniform topographic and hydrologic characteristics at the grid scale. More importantly for this study, CLSM is one of the few modern, physically based, distributed land surface models that simulate unconfined groundwater storage variations, whereas most have a lower boundary in the unsaturated zone, typically 2–3 m below the land surface. Groundwater must be simulated in order to generate terrestrial water storage variations that are analogous to those measured by GRACE. The catchment delineation (demonstrated for the East Coast and California Basin in Figure 1) is based on a 30 arc sec ($\sim$1 km) digital elevation model from the USGS [*Verdin and Verdin*, 1999] and results in an average catchment size of 3640 km^2 in North America. For reasons of computational efficiency, catchments are further divided into "tile" units, defined by the intersection of catchments with the overlying atmospheric grid (Figure 1), and model simulations are performed separately for each "tile."

[13] CLSM simulates subgrid hydrological processes (root zone soil moisture distributions) based on each catchment's topographical statistics. With this information the catchment is partitioned into dynamically varying areas of different hydrological regimes ("saturated," "transpiring," and "wilting") each governed by distinct evaporation and runoff parameterizations [*Koster et al.*, 2000]. These distinct regimes represent regions where the root zone is fully saturated, the unsaturated root zone moistures lie above the vegetation-specific wilting point, and transpiration is shut off completely, respectively. CLSM simulates the moisture content of all layers down to and within the saturated zone as an equilibrium catchment deficit (the average depth of water needed to bring all of the soil throughout the catchment to saturation). The model imposes a maximum catchment deficit based on a standardized global data set of soil profile depths (depth to bedrock) compiled from the FAO/UNESCO Soil Map of the World [*Webb et al.*, 1991]. The notion of bedrock being a perfectly fixed and unbroken lower boundary condition is a modeling convenience that would limit the value of CLSM output for certain local hydrogeological applications, but it has little bearing on the primary goal of this study, which is to use model-assimilated groundwater levels to assess drought at scales of tens of kilometers and larger. However, the maximum catchment deficit does have implications for the model's ability to simulate severe drought conditions as described in section 2.5 and discussed further in section 3. Temporal changes in the catchment deficit represent column-integrated variations in

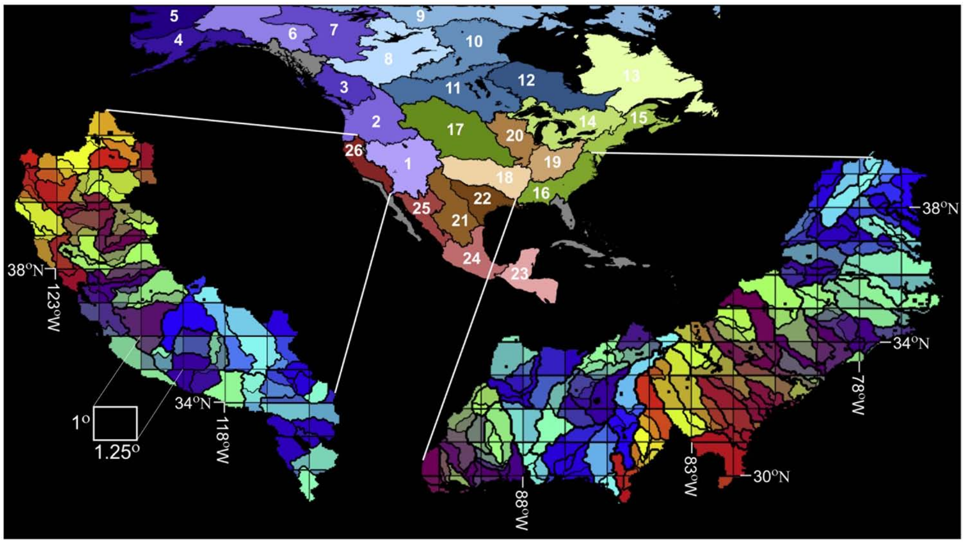

Figure 1. Map of hydrologically defined basins for North America used for extracting GRACE-derived terrestrial water storage anomalies. The catchment delineation used by the Catchment Land Surface Model (CLSM) is shown for two selected basins. The average size of the catchments in North America is 3640 km^2. The overlaying grid represents the resolution of the atmospheric forcing data. The intersection of a catchment with the overlying atmospheric grid defines the "tile," which is the fundamental CLSM model unit. Within the U.S. domain the basin numbers denote the following named basins: 1, Great Basin and Colorado; 2, Columbia; 14, Great Lakes; 15, Upper East Coast; 16, East Coast; 17, Missouri; 18, Arkansas-Red and Lower Mississippi; 19, Ohio; 20, Upper Mississippi; 22, Gulf; 26, California.

subsurface water storage. Two additional CLSM prognostic variables, root zone excess and surface excess, provide a representation of nonequilibrium profile conditions near the surface, whereas changes in snow water storage are modeled using a state-of-the-art three-layer snow physics scheme [*Stieglitz et al.*, 2001]. Like most distributed land surface models, CLSM does not simulate lateral moisture flows between spatial elements (catchments or tiles). At the spatial and temporal scales relevant here, lateral flows of groundwater and soil moisture are negligible compared with the vertical fluxes. Further, for the contiguous United States, contributions from surface water and biomass to terrestrial water storage variability occur at or below the uncertainty level of GRACE observations [*Rodell and Famiglietti*, 2001; *Rodell et al.*, 2005a], and the CLSM water reservoirs thus enable estimation of TWS anomalies (deviations from the mean) which are directly analogous to the GRACE observations. Groundwater storage can be calculated by subtracting the total root zone moisture content (in equivalent heights of water) and snow water equivalent from the total column integrated water storage.

2.3. GRACE DAS

[14] A detailed description of the GRACE DAS is provided by *Zaitchik et al.* [2008, Figure 5] and only a brief overview is given here. GRACE DAS aims to synthesize the advantages of GRACE TWS observations and CLSM soil moisture and groundwater estimates to generate enhanced estimates of individual TWS components, informed by GRACE and with the high spatial and temporal resolutions and timeliness of the model (and its other input data), based on an ensemble Kalman smoother (EnKS) algorithm. The modeled moisture fields are corrected toward the observational GRACE estimate to a degree determined by the relative uncertainty in the model and the observations. The EnKS downscales these basin scale corrections to assimilation increments which are applied to the numerous subbasin model elements (the "tiles" shown in Figure 1) on the basis of modeled error correlations between the basin-scale TWS estimates and the tile space components of TWS, including soil moisture and groundwater. These error correlations are modeled in CLSM through an ensemble-based approach using 20 members by applying perturbations to select surface meteorological forcing fields and CLSM prognostic variables [*Zaitchik et al.*, 2008; *Liu et al.*, 2011], which is a common approach in land data assimilation that accounts for uncertainty in model parameters as well as model structure. Specifically, spatially correlated, temporally correlated, and cross-correlated random fields are applied to precipitation (multiplicative perturbations with standard deviation $s = 0.5$), shortwave radiation (additive with $s = 0.3$ W m^{-2}), and longwave radiation (additive with $s = 50$ W m^{-2}) at every 3 h forcing time step. Cross correlations (ρ) were imposed between perturbations to precipitation and longwave radiation ($\rho = 0.5$), precipitation and shortwave radiation ($\rho = -0.8$), and longwave and shortwave radiation ($\rho = -0.5$). Furthermore, the CLSM prognostic variables catchment deficit and surface excess were perturbed at every 20 min model time step with standard deviations of 0.05 and 0.02 mm, respectively. In all cases, a stationary horizontal error correlation with *e*-folding scale of 2° (~200 km) was assumed. Covariance localization was applied to the calculation of horizontal error correlations to prevent spurious long-range correlations in the forecast ensemble [*Reichle and Koster*, 2003]. We used temporal correlation (*e*-folding)

scales of 96 h for forcing fields and 24 h for model prognostic fields. The temporal and horizontal error correlation scales were estimated on the basis of previous experience with CLSM [*Reichle and Koster*, 2003; *Zaitchik et al.*, 2008].

[15] In GRACE DAS, the tiles within a basin are thus coupled in the data assimilation system through the horizontally correlated errors. The system therefore implicitly accounts for unmodeled fluxes across catchment boundaries, spatially correlated errors in parameter fields, and spatially correlated errors in meteorological forcing data [*Reichle and Koster*, 2003]. Errors in the GRACE estimates result from measurement and processing errors from multiple sources. We assumed a conservative RMS error estimate of 20 mm for GRACE basin averages [*Wahr et al.*, 2006]. We also performed a second assimilation run using a GRACE error of only 10 mm in order to assess the effect of GRACE uncertainty estimates on data assimilation skill.

2.4. Forcing Data and Model Initialization

[16] The CLSM was run for the period from August 2002 to July 2009 for the North American Domain (Figure 1) using a combination of forcing data sets as described below. In addition, a long-term (1948–2009) CLSM simulation was performed to establish a reference for creating the drought indicator percentiles in a manner consistent with the U.S. and North American Drought Monitors (section 2.6).

[17] For the long-term run, the CLSM was forced using globally consistent, near-surface meteorological data from Princeton's Global Meteorological Forcing Data set covering the period from 1948 to 2006 [*Sheffield et al.*, 2006]. The Princeton data set is 3-hourly with a 1° grid resolution. It was produced by combining observation-based data sets with a reanalysis product. The open-loop (OL; no data assimilation) and data assimilation (DA) runs for the August 2002 to July 2009 period used a combination of observation-based forcing data sets from the North American and Global Land Data Assimilation Systems (NLDAS and GLDAS) (http://ldas.gsfc.nasa.gov/). We used the NLDAS-2 data set (1/8th° and hourly) over central North America (−125°W to −67°W, 25°N to 53°N) which is constructed from gauge-based observed precipitation, bias-corrected shortwave radiation and surface meteorology analysis fields of the NCEP North American Regional Reanalysis (NARR). GLDAS forcing (1/4° and 3 hourly) was used for the remainder of North America consisting of NOAA Climate Prediction Center's Merged Analysis of Precipitation product [*Xie and Arkin*, 1997], radiation fields from the Air Force Weather Agency, and atmospheric reanalysis fields from NCEP's Global Data Assimilation System (GDAS) [*Kleist et al.*, 2008]. The NLDAS-2 and GLDAS data sources were fused and resampled to produce a hybrid 3-hourly LDAS forcing data set at 1° × 1.25° spatial resolution.

[18] The CLSM was initialized to equilibrium conditions by using climatological average states from the model for the precise time of year of initialization. This approach has been determined to be preferable to the traditional single-year spin-up method [*Rodell et al.*, 2005b]. The mean state fields required to initialize the Princeton run (1948–2006) were generated by looping 10 times through 15 years (1948–1963) of Princeton forcing and computing the average of the output from the last complete loop for the precise time of year of the start of the experimental period (1 January 1948). The same approach was applied to the Hybrid run (2002–2009) by looping 10 times through 8 years (August 2002 to July 2009) of Hybrid LDAS forcing.

2.5. Increasing the Water Storage Capacity of the CLSM

[19] Analysis of initial GRACE DAS output revealed that simulated TWS did not dry as much as GRACE observations indicated during extended dry periods, because the dry anomaly observed by GRACE exceeded CLSM's maximum possible catchment deficit. For the purpose of drought monitoring it is crucial that the CLSM catchment deficit has a sufficiently broad range of variability to represent severe drought conditions and thus accommodate the driest TWS anomalies that could be observed by GRACE and assimilated. Figure 2 maps the probability of the unaltered CLSM reaching the tile-specific dry limits in TWS over an 8 year simulation period (2002–2009). The dry limits were determined by forcing CLSM with zero precipitation over the entire study period. The issue is particularly evident in the western United States and Mexico where modeled TWS frequently approaches the lower limit defined by the tile-specific thresholds. The time series records for selected sites in California and Mexico illustrate that hitting the dry limit can be a nearly annual occurrence, which would cause an abrupt shut down of soil moisture and groundwater depletion as the model has no more water left to remove (Figure 2). The upshot is that the severity of simulated drought events is effectively capped. To remedy this situation we increased the water storage capacity of CLSM by increasing the depth to bedrock parameter uniformly by 2 m. Note that such an increase is not inconsistent with the bedrock depth data used in the original CLSM. The soil profile depths in the global data set represent minimum possible values because the profile descriptions did not always extend to the subsurface bedrock [*Webb et al.*, 1991]. This adjustment resulted in more realistic simulated dry down that was not constrained by an artificial dry limit during the 8 year period (Figure 2). Regions not initially impacted by a too narrow dynamic TWS range exhibited only minor changes in the TWS variability as a result of the bedrock depth increase (Figure 2). Additional sensitivity experiments indicated that this modification had minimal impacts on simulated fluxes and hydrological states in these regions (not shown). The effect of the depth to bedrock increase on the model's ability to predict the magnitude and seasonality of observed TWS and groundwater storage is discussed in sections 3.1 and 3.2.

2.6. Drought Monitoring

[20] A key objective of this study was to deliver supplemental drought indicators, which have been informed by GRACE TWS data, for integration into the U.S. and North American Drought Monitors. The U.S. Drought Monitor (USDM) was conceived in 1999 through the cooperation of the U.S. Department of Agriculture (USDA) and the National Drought Mitigation Center (NDMC) with the goal of centralizing the drought monitoring activities conducted by federal, state, and academic entities in the United States [*Svoboda et al.*, 2002]. Similarly, the North American Drought Monitor (NADM) has been providing integrated assessments of drought throughout most of Canada, Mexico, and the United States on a monthly basis since 2003

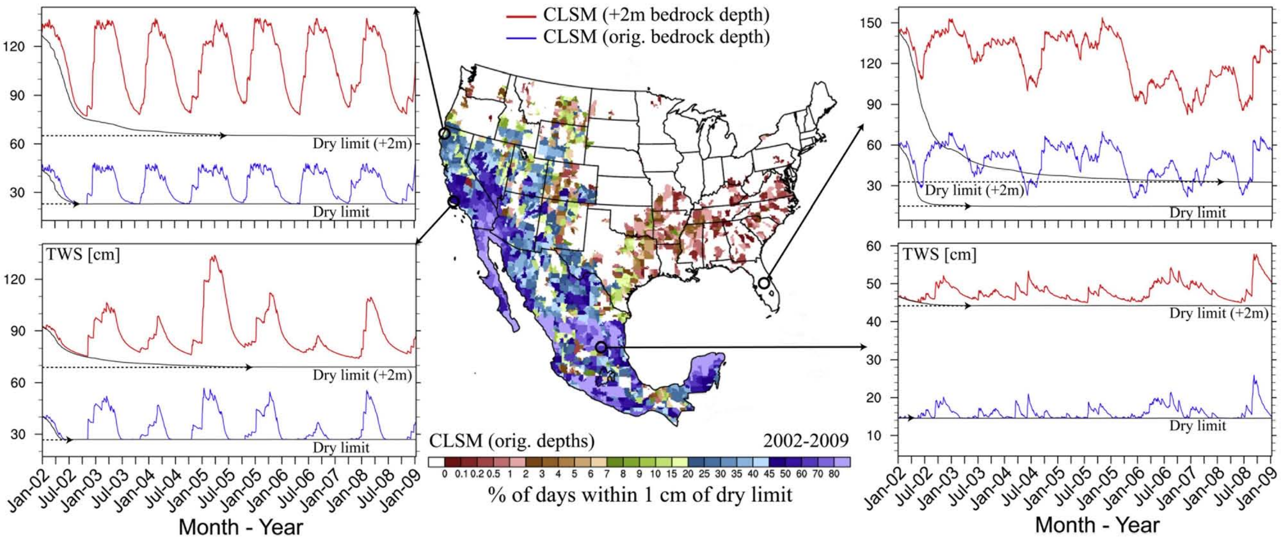

Figure 2. The map of the U.S. and Mexican domain depicts the percentage of days over the 2002–2009 period when the terrestrial water storage (TWS) simulations by CLSM were at or near (within 1 cm) the tile-specific dry limits. The dry limits were established by forcing the CLSM with zero precipitation over the simulation period (indicated by the solid black lines with arrows). The accompanying time series records serve to illustrate how the TWS often drops down to the tile-specific thresholds during extreme drying events, causing an abrupt shut down of the drying trend as the model has no more water left to remove. This issue is resolved by increasing the depth to bedrock uniformly by 2 m as shown by the top (red) curves.

[*Lawrimore et al.*, 2002]. USDM maps are produced weekly, led by a rotating group of authors from both federal agencies and academia. The authors rely on a suite of short- and long-term objective indicators as well as subjective input from a network of water and climate experts at the local and regional level.

[21] The purpose of the drought monitors is to provide timely, understandable, consensus information on water supply and drought for decision makers, stakeholders, and the general public. Both classify drought severity into four major categories (i.e., D4, D3, D2, and D1) with a fifth category (D0) depicting "abnormally dry" conditions (Table 1). Each category is associated with its probability of occurrence, expressed as a percentile, on the basis of past data. For instance, D3 (extreme) drought conditions have occurred at any given location 2%–5% of the time (second to fifth percentile). A 1932–2001 data record of drought indicators is used to estimate the frequency of occurrence of a given drought category for the location and time of year (month) in question [*Svoboda et al.*, 2002].

[22] As a starting point, the USDM authors use the previous week's map and short- and long-term objective blends (defined in section 1). The authors then subjectively incorporate streamflow and other in situ observations, satellite-based drought indicators including vegetation greenness, and reports from local agencies and stakeholders to draw the final map. None of the previously available indicators directly represented deep soil moisture and groundwater storage conditions. Thus our overarching hypothesis, which is still being evaluated, is that drought conditions can be described more comprehensively and more objectively by incorporating GRACE-based soil moisture and groundwater information into the drought monitor production process.

Table 1. Categories of Drought Magnitude Used in the Drought Monitors[a]

Category	Drought Severity Level	Percentile
D0	Abnormally dry	20–30
D1	Drought, moderate	10–20
D2	Drought, severe	5–10
D3	Drought, extreme	2–5
D4	Drought, exceptional	≤2

[a]Percentile thresholds and are depicted for each severity level.

2.7. GRACE-Based Drought Indicators

[23] Three CLSM prognostic moisture variables were translated into drought indicators (DI): surface soil moisture (sfsm), root zone soil moisture (rtzsm) and groundwater storage (gws). The integration of these fields into the drought monitors requires conversion to percentiles using an approach consistent with the drought monitor process (see above). Accordingly, 8 years of GRACE-assimilated fields of soil moisture and groundwater were converted into drought indicator percentiles using a climatology of soil moisture and groundwater based on a 62 year CLSM simulation (1948–2009). Calculating meaningful, current drought indicator percentiles requires a consistent long-term data record that extends to present time. The long-term simulation was forced by two different meteorological forcing sources (section 2.4), which caused a discontinuity in the output moisture fields. Figure 3a depicts differences in the 2002–2006 mean of Princeton and LDAS-forced surface soil moisture output for the North American domain. Biases are evident in the U.S. Southeast, in parts of Canada, and in Mexico. The Princeton-forced output fields are considerably wetter, with soil moisture differences occasionally exceeding 0.11 m^3 m^{-3}. The map also demonstrates that biases are

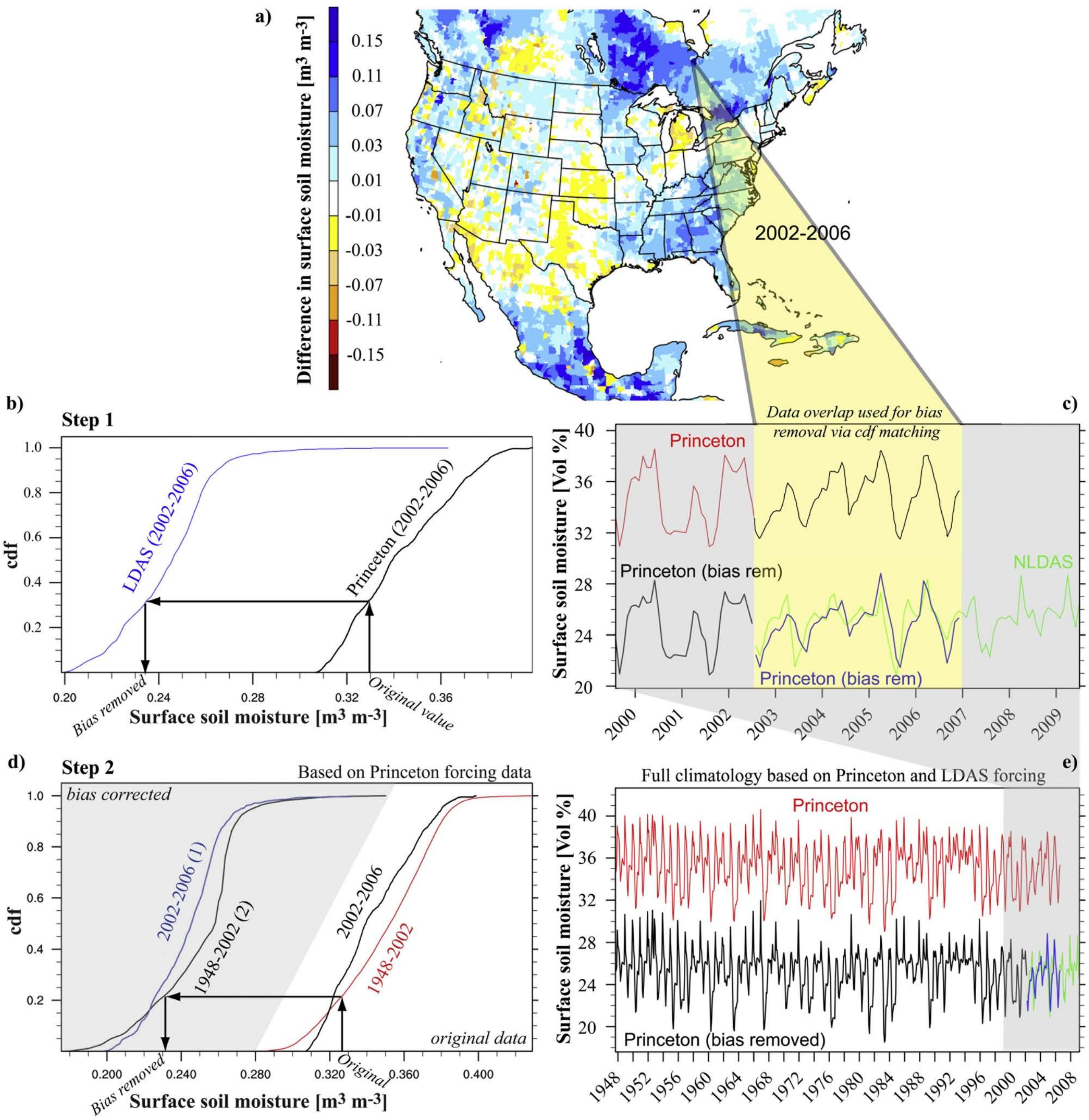

Figure 3. Bias adjustment scheme for combining CLSM moisture output based on two different meteorological forcing sources (LDAS and Princeton) and to ensure consistency of full climatology over the 1948–2009 simulation period. (a) Differences in the mean (2002–2006) of Princeton and LDAS surface soil moisture (positive values indicate that the Princeton-based soil moisture is larger). (b) The CDF of the 2002–2006 Princeton-forced moisture fields matched to that of the LDAS-forced moisture fields as illustrated for a sample catchment tile in Canada. (c) Resulting time series of the original and bias-adjusted soil moisture data. (d) The bias-corrected Princeton CDF (2002–2006) resulting from step 1 (blue) used in association with the original Princeton short-term (black) and long-term (red) CDFs to generate a new long-term CDF (black) representative of the full bias-corrected Princeton data set. (e) Consistency of the bias-corrected soil moisture values across the entire climatology period.

nonuniform and exhibit complex spatial patterns. While the Princeton-forced fields are biased in relation to the LDAS-forced fields, we make no assumptions about biases with respect to real-world moisture conditions.

[24] Therefore, a statistical adjustment was applied to the earlier output (1948–2002) on the basis of parallel simulations over the overlapping period of the Princeton and LDAS forcing data sets (2002–2006; Figure 3). Following

Reichle and Koster [2004], the cumulative distribution function (CDF) of the 2002–2006 Princeton-forced moisture fields was matched to that of the LDAS-forced moisture fields, as illustrated in Figure 3b for a single catchment tile in Canada. Time series of the original and bias adjusted soil moisture data are shown in Figure 3c. Next, the resultant bias-corrected Princeton-forced CDF (2002–2006) was used in association with the original short-term (2002–2006) and long-term (1948–2002) Princeton CDFs to adjust the CDF of the full Princeton-forced data set ($CDF_{bc,1948–2002}$; Figure 3d) according to

$$CDF_{bc,1948-2002} = (CDF_{1948-2002} - CDF_{2002-2006}) + CDF_{bc,2002-2006}. \quad (1)$$

[25] The full Princeton forced data set was then bias corrected using $CDF_{bc,1948–2002}$ and $CDF_{1948–2002}$ as illustrated in Figure 3d. This approach was applied to each catchment tile unit of the North American domain. The advantage of the adopted bias adjustment technique is that it preserves the differences observed between the original 2002–2006 and 1948–2002 Princeton CDFs (Figure 3d). The consistency of the bias-corrected soil moisture values across the entire climatology period (1948–2009) is demonstrated in Figure 3e.

[26] Finally, modeled surface soil moisture, root zone soil moisture, and groundwater storage were converted to percentiles by ranking the values against the bias-corrected long-term climatology. Location (i.e., tile) and time-specific (i.e., for each month) CDFs were generated for each moisture component on the basis of the full bias-corrected historic data set (1948–2009) and used to convert absolute moisture values into percentiles corresponding to the drought severity classification adopted by the drought monitors (Table 1).

2.8. Evaluation Data Sets

[27] Moisture output from GRACE DAS was evaluated against various independent data sets comprising runoff-calibrated model simulations of TWS and in situ network observations of groundwater and soil moisture as described below.

2.8.1. VIC Model Output

[28] Independent model estimates of TWS for the California (Figure 1, basin 26), Columbia (2), and Missouri (17) basins were obtained from a simulation of the variable infiltration capacity (VIC) macroscale hydrology model [*Liang et al.*, 1994] as configured and run by *Gao et al.* [2010]. While this particular VIC simulation made use of the same recent-period forcing data, from NLDAS, as did our CLSM simulations (section 2.4), the *Gao et al.* [2010] VIC simulation was calibrated using observed streamflow across much of the continental United States. The VIC TWS estimates were derived by adding together soil moisture at three levels and snow water equivalent. Groundwater is not explicitly represented in VIC, but *Gao et al.* [2010] found that the dynamic range of TWS in VIC was significantly larger than that in the GRACE Tellus product in both the California and Columbia basins. Motivated in part by this result, scaling coefficients (described in section 2.1) were developed and are now provided with the GRACE Tellus product to counteract the signal attenuation effects of GRACE data processing. Still, VIC's dynamic ranges of TWS in these two basins are similar to those of scaled GRACE Tellus (see section 3.1). Thus it is likely that overly large modeled soil moisture variations compensate for the lack of groundwater in VIC. Time series of VIC monthly mean TWS anomalies over the 2003–2007 period were used in our evaluation.

2.8.2. Groundwater Observations

[29] Groundwater storage variations were derived from in situ groundwater level measurements from 239 monitoring wells across the United States after rigorous examination of data quality and suitability (Figure 4a). Daily or monthly well measurements were acquired from the USGS groundwater watch (http://groundwaterwatch.usgs.gov/), the Texas Water Development Board (http://www.twdb.state.tx.us/gwrd/waterlevels/waterlevels.html), and the Illinois State Water Survey network (http://www.isws.illinois.edu/warm/sgwdata/wells.aspx) over the August 2002 to July 2009 GRACE period, as available. Many locations were excluded from the analysis because of data gaps and we only used data from wells that captured the seasonal cycle (typically near-surface unconfined or semiconfined aquifers, as inferred from available metadata for each well and published reports

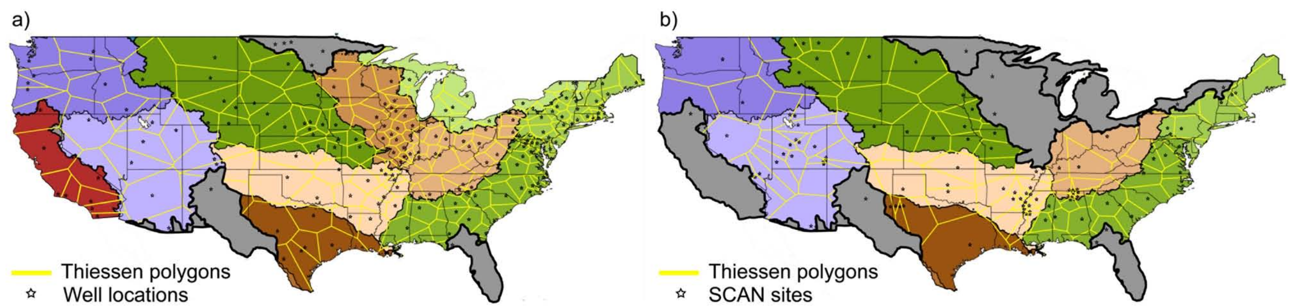

Figure 4. (a) Spatial distribution of the 239 monitoring wells used as the basis for validating model simulations of groundwater storage. (b) Spatial distribution of the Soil Climate Analysis Network (SCAN) sites providing time series of surface soil moisture measurements. Thiessen polygons were generated to compute basin average measured time series of groundwater and surface soil moisture. The basin delineations are as in Figure 1. Basins in gray were masked out from the analysis because of poor network coverage.

on the local stratigraphy) and that were not directly affected by pumping (as indicated in the metadata), meaning that the well was neither itself pumped nor in the immediate vicinity of a well being pumped at the times of observation. At each well a representative value of the specific yield (Sy) was determined in order to convert well water level measurements to equivalent heights of stored water. Estimating appropriate specific yield values is important because uncertainties in Sy will obscure comparisons between GRACE-based and in situ groundwater estimates; even a small change in Sy can change the amplitude of the computed groundwater fluctuations significantly [*Rodell et al.*, 2007]. Therefore we put considerable effort into the selection of Sy values, using any available metadata on the name and/or material composition of the geologic formation and an extensive review of reports published by the USGS. Resulting estimates ranged from 0.01 to 0.35 with a mean of 0.13.

[30] Daily groundwater storage anomaly time series were then generated for each well site. The Thiessen polygon method was applied to subdivide the basins and determine the area weight of each site toward the regional basin average (Figure 4a). Resulting area-weighted basin-averaged groundwater anomaly time series were used to evaluate groundwater storage fluctuations estimated by GRACE DAS. Spatial undersampling of groundwater levels in regions where appropriate well records are sparse, particularly in the western United States (Figure 4a), was another source of uncertainty. Therefore, certain areas were completely excluded from the analysis because of poor data coverage.

2.8.3. Soil Moisture Observations

[31] Time series of surface and root zone soil moisture were obtained from Soil Climate Analysis Network (SCAN) sites (Figure 4b) operated by the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) through the National Water and Climate Center (http://www.wcc.nrcs.usda.gov/scan/). Hourly soil moisture measurements at 5, 10, 20 and 50 cm depths were extracted from each site over the 2002–2009 study period and converted into daily time series, disregarding values recorded during freezing soil temperatures. The 5 cm SCAN observation was assumed representative of surface soil moisture whereas a layer-depth-weighted average of all 4 levels was used for deriving root zone soil moisture. Basin-averaged time series were generated using Thiessen polygons (Figure 4b). The California (26), Rio Grande (21), Upper Mississippi (20), and Great Lakes (14) basins were excluded from the analysis because of poor SCAN network coverage. The area-weighted basin-averaged soil moisture (sm) time series were normalized using the maximum and minimum for each basin (i.e., $[sm_i - sm_{min}]/[sm_{max} - sm_{min}]$) before comparing them to model simulations of surface (0 – 2 cm) and root zone (0 – 100 cm) soil moisture normalized in a similar manner.

3. Results

3.1. Terrestrial Water Storage

[32] There is generally a good correspondence between the open-loop simulated seasonal TWS cycle and that observed by GRACE for the selected basins, with characteristic wintertime peaks in TWS followed by a summer trough (Figure 5). As expected, the data assimilation resulted in time series intermediate between the open-loop (with bedrock depth increased by 2 m; OL2) simulation and the GRACE observations (Figure 5). The CLSM open-loop and VIC time series agree nicely for all three basins as evidenced by Pearson's correlation coefficients (r) ranging from 0.93 to 0.97 (OL) and 0.90 to 0.98 (OL2; Table 2), which is expected given the similarity of the applied meteorological forcing. Interestingly, for the California (CA) and Missouri (MO) basins the CLSM output based on the original bedrock depths (OL) provides a better fit to the VIC data. One plausible explanation may be that VIC is also impacted by limited water storage capacity to some extent. This hypothesis is supported by VIC's simulated

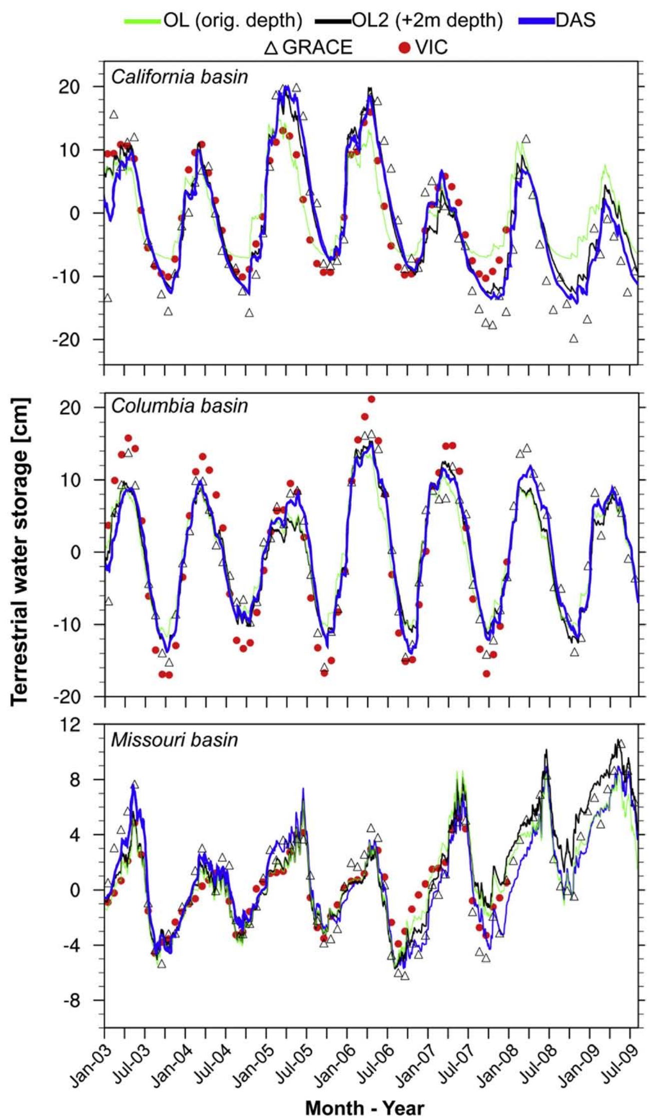

Figure 5. Time series plots of TWS for the California, Columbia, and Missouri basins comparing monthly variable infiltration capacity (VIC) model output with open-loop simulations based on the original bedrock depths (OL) and increased bedrock depths (OL2), data assimilation output (DAS), and GRACE TWS observations.

Table 2. Pearson's Correlation Coefficient r and Root-Mean-Square (RMS) Difference Resulting From Comparing Terrestrial Water Storage Simulations From VIC and CLSM for the California, Columbia, and Missouri Basins[a]

	California		Columbia		Missouri	
VIC-CLSM	r	RMS (mm)	r	RMS (mm)	r	RMS (mm)
VIC-CLSM (OL)	0.96	2.68	0.97	4.62	0.93	1.06
VIC-CLSM (OL2)	0.93	3.29	0.98	3.94	0.90	1.22
VIC-CLSM (DAS)	0.89	4.06	0.97	3.93	0.88	1.50
VIC-GRACE	0.82	6.44	0.96	3.55	0.87	1.76

[a]See Figure 5. Results are shown for open-loop simulations with original bedrock depths (OL) and with bedrock depths uniformly increased by 2 m (OL2) as described in section 2.5. The data assimilation (DAS) results were derived using a GRACE observation error of 20 mm. The agreement with the GRACE terrestrial water storage estimates is also shown. CLSM, Catchment Land Surface Model; VIC, variable infiltration capacity macroscale hydrology model.

TWS for the CA basin, which hit almost the exact same minimum every year (−11 cm), suggesting that there is no more water to remove. CLSM simulations with the increased bedrock depths (OL2) exhibited more realistic variability in the annual minima. GRACE TWS observations corroborate the idea that the extent of TWS depletion varies from year to year in the CA basin (Figure 5).

[33] The average seasonal amplitude of TWS from the OL simulations in CA (16.3 cm) is significantly smaller than that indicated by the GRACE observations (27.9 cm). The open-loop simulations with increased bedrock depths (OL2) better approximate the dynamic range of the GRACE observations (Figure 5). The effect of data assimilation (DAS) is generally minimal because the OL2 and GRACE results already agree fairly well (Figure 5). However, one might expect a larger correction downward toward the GRACE observations during the summer/fall minimum. A likely explanation is that the data assimilation scheme was unable to dry the land surface much more because it relies too much on perturbations to the precipitation forcing field to generate a spread in the ensemble, and when precipitation is already low that spread is necessarily small. (In future experiments we will explore increasing the direct perturbation of the prognostic water storage variables to address this issue.) The CLSM specific dry limit (Figure 2) is not reached in the Columbia (CB) basin, and as a result there is little difference between the OL and OL2 simulations (Figure 5).

[34] For the CB basin, however, the VIC simulations are characterized by a larger amplitude (31.4 cm compared to 20.9 cm for OL2), which is more in line with the GRACE observations (27 cm). Accordingly, GRACE data assimilation increases the amplitude of the TWS simulations from 20.9 to 21.5 cm on average, resulting in reduced RMS errors when compared against the VIC simulations (Table 2).

[35] Significantly larger discrepancies between VIC and GRACE TWS were reported by *Tang et al.* [2010] and *Gao et al.* [2010] for subbasins within the CA and CB basins. Spatial "leakage" of the gravity signal was implicated as a major contributor to the apparent underestimation of seasonal TWS amplitude by GRACE relative to VIC. This leakage of the gravity signal across the boundaries of a region of interest is analogous to blurred vision, and it results from GRACE's low spatial resolution and the numerical processing required to extract a time series for a particular region of interest from the global GRACE gravity solutions [*Wahr et al.*, 1998]. *Tang et al.* [2010] and *Gao et al.* [2010] used a GRACE product that was not optimized for land applications. In this study, scaling methods were applied to restore some of the power in the GRACE signal attenuated by data processing (section 2.1). Using a GRACE product optimized for hydrology evidently reduces the biases between the modeled and observed estimates to within acceptable limits for the displayed basins (Figure 5).

3.2. Groundwater Storage

[36] One of the advantages of the GRACE data assimilation system is its ability to decompose the vertically integrated GRACE TWS signal into groundwater, soil moisture, and snow, which individually are more valuable for scientific applications. Figure 6 compares open-loop (OL and OL2) and data assimilation (DAS) simulations of groundwater storage with well observations that were resampled to basin averages using Thiessen polygons as detailed in section 2.8.2. The time series are expressed as anomalies with respect to the period mean (2002 – 2009). The monthly GRACE TWS data are also plotted to help visualize the impact of the data assimilation. Wide discrepancies exist between the OL and OL2 simulations particularly for the Great Basin and Colorado basin and the East Coast basin (Figure 6). Increasing the bedrock depth provides the potential for a wider dynamic water storage range (i.e., larger amplitude). Since more water is allowed to enter the system there is the opportunity for enhanced subsequent drying as clearly seen in the East Coast basin (Figure 6). The measured groundwater time series provide some justification for the larger seasonal amplitudes. While the OL2 simulations do tend to overestimate the anomalies during dry and wet events relative to the well observations (Figure 6), GRACE data assimilation generally succeeds in adjusting the groundwater storage estimates to better match the measurements. However, as described in section 2.8.2, the magnitude of the measured groundwater storage fluctuations is uncertain given the difficulties associated with assigning representative specific yields for each well site.

[37] Figure 7a evaluates skill improvement associated with the bedrock depth adjustment defined as skill (r) of OL2 model output minus skill (r) of OL model estimates. Here the skill (r) represents the correlation between the time series of the open-loop estimates (expressed as anomalies relative to their 2002–2009 climatology) and the anomaly time series of the basin-averaged groundwater measurements. Skill improvements are evident in the Great Basin and Colorado basin (1) and the California basin (26), which frequently encounter water storage simulations within the dry limit of the model when using the original bedrock depths (section 2.5). Improvements are also seen in the Arkansas and Lower Mississippi basin (18), the Upper Mississippi basin (20) and the upper part of the East Coast basin (15). Most of the eastern United States is characterized by a decrease in correlation as a result of increasing the bedrock depths, but decreases are only statistically significant (5% level) in the Ohio basin (19).

[38] The skill associated with assimilating the GRACE TWS data is assessed in Figure 7b, which depicts the skill

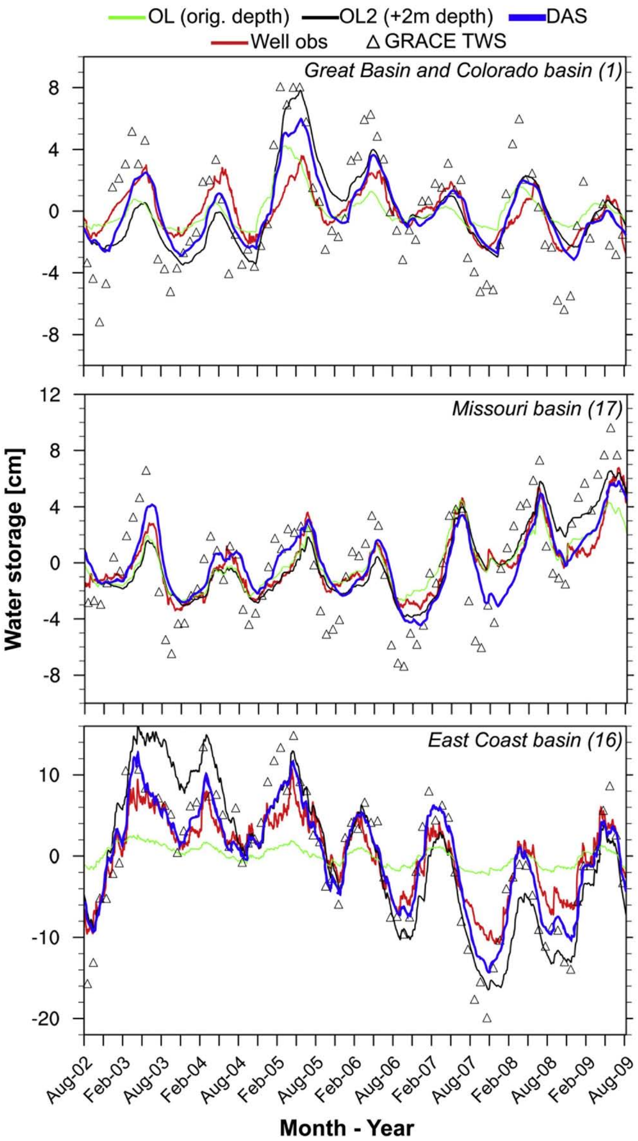

Figure 6. Time series plots of groundwater storage for the Great Basin and Colorado basin, Missouri basin, and East Coast basin comparing daily well groundwater measurements with open-loop simulations based on the original (OL) and increased bedrock depths (OL2) and data assimilation output (DAS). GRACE TWS observations are overplotted to better visualize the impact of the data assimilation.

of the data assimilation (DAS) simulations minus the skill of the OL2 results as in Figure 7a. Table 3 lists the associated statistics including r and RMS errors for two data assimilation runs with GRACE observation errors of 20 mm and 10 mm. Basins that experienced statistically significant improvements in skill in Figure 7a (basin 1, 18 and 15) also see increases in data assimilation skill (Figure 7b), which are all statistically significant (Table 3) suggesting that the GRACE data contain some independent information that the assimilation algorithm is able to translate into superior model estimates. The value of the GRACE data assimilation is also clearly evident for the Ohio basin (19) where r increases from 0.86 to 0.94 and the RMS error decreases from 4.94 to 2.34 (Table 3), which brings the correlation back to the level of the OL simulations (Figure 7a). This may suggest that the bedrock depth adjustment is not warranted in this area. Similar tendencies are evident in the East Coast basin (16) and Great Lakes basin (14), where TWS only rarely exceeds the maximum possible catchment deficit of CLSM using the original bedrock depths. However while the increased bedrock depths may result in a too wide dynamic water storage range in some parts of the eastern United States, it is hard to find justification for using the original bedrock depths when comparing groundwater time series for the East Coast basin (Figure 6). Here the data assimilation results in a time series intermediate between the OL2 and OL simulations, which appears to provide a better approximation to the observed values. For drought monitoring, sufficient simulated water storage capacity is critical, as it allows the model to distinguish the severity of different drought events. Further, in this particular application it ensures that dry anomalies observed by GRACE do not exceed the maximum possible catchment deficit of CLSM (i.e., the dry limit) [*Zaitchik et al.*, 2008] so that the data assimilation scheme can work as intended.

[39] Negative skill scores ($r_{DAS} - r_{OL2}$), with statistical significance at the 5% level, occur in the Missouri basin (basin 17) and California basin (basin 26; Figure 7b). In the Missouri basin discrepancies between data assimilation results and the groundwater measurements is mainly the result of a drying trend in the fall of 2007 indicated by the GRACE observations but not reported in the well records (Figure 6). According to the well records 2006 experienced the largest dry anomaly whereas 2007 and 2008 were characterized by similar and much wetter conditions. According to U.S. Drought Monitor maps (not shown) the entire basin was in a D0 to D3 drought (see Table 1) in 2006 (August–December); in 2007 (September–December) D0 to D2 drought conditions were confined to the western section of the basin (excluding almost entirely the states of Nebraska, Iowa, Missouri and Kansas); and in 2008 the drought abated across most of the basin. The difference in drought conditions between 2007 and 2008 suggested by the U.S. Drought Monitor conflicts with the observed well time series record. Many of the wells that were used to calculate the basin-averaged anomaly time series are concentrated in the southeastern part of the Missouri basin (Figure 4) within states that were largely unaffected by drought conditions during 2007, which may have impacted the representation of basin-wide groundwater storage, despite the use of Thiessen polygons to area weight the well data. Spatial undersampling may also be a plausible explanation for the discrepancy between the data assimilation results and the groundwater measurements in the CA basin as the majority of the wells are located in the southwestern section with only a couple of wells located in the interior of the state (Figure 4) where climatic conditions are significantly different.

[40] The data assimilation results displayed in Figures 6 and 7 were derived using a GRACE observation error of 20 mm. This is a fairly conservative estimate and simulation runs assuming a 10 mm RMS error were executed for comparison. The effect of using a reduced observation error is generally minor, with correlations decreasing slightly and RMS errors remaining largely unchanged (Table 3).

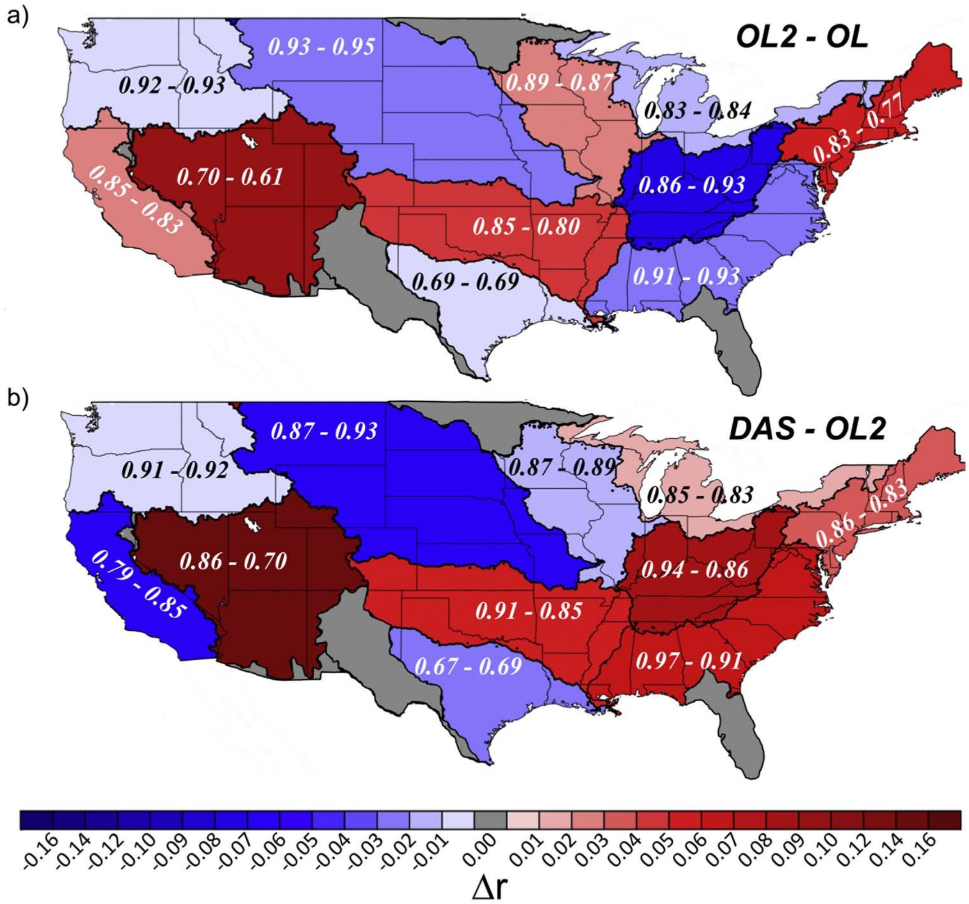

Figure 7. (a) Differences in time series correlation coefficients (r) between the OL2 and OL model simulations of groundwater storage. The skill (r) represents the correlation between measured (wells) and model-simulated anomaly time series of basin-averaged groundwater. Reddish colors indicate that model skill is improved as a result of increasing the bedrock depths by 2m. (b) Differences in skill between the data assimilation (DAS) and open-loop (OL2) simulations. Reddish colors indicate that model skill is improved as a result of assimilating GRACE TWS data. The actual correlation coefficients are provided for each basin and simulation type directly on the plots.

3.3. Soil Moisture

[41] Figure 8 showcases weekly time series of surface soil moisture resulting from open-loop (OL2) and data assimilation runs and basin-averaged SCAN in situ observations for the Ohio (19) and East Coast (16) basins. Generally, there's a good correspondence between the model estimates and surface soil moisture measurements when expressed in normalized units (see section 2.8.3). Both basins are characterized by positive skill scores for data assimilation ($r_{DAS} - r_{OL2}$) that are statistically significant for surface soil moisture (sfsm) as well as root zone soil moisture (rtzsm) (Table 4). Basins in the midwestern and western United States generally had insignificant increases/decreases in correlation. Also for soil moisture, the effect of reducing the GRACE observation error to 10 mm had a minor effect on the correlation statistics (not shown). While spatial undersampling due to the limited number of SCAN sites (Figure 4b) may impact the comparability between basin-averaged soil moisture measurements and model estimates, the positive skill scores, particularly in the eastern United States, suggest that GRACE data assimilation is also valuable for soil moisture. However, as expected, the assimilation of monthly GRACE TWS data appears to have a larger impact on groundwater, which varies slowly, compared to soil moisture, which responds more quickly to short-term variations in atmospheric forcing.

3.4. Drought Indicators

[42] Drought Indicators (DI) for potential integration into the U.S. and North American Drought Monitors were generated by converting moisture fields from GRACE DAS into DI percentiles using a 62 year climatology of CLSM simulated soil moisture and groundwater (section 2.7). Figure 9 illustrates the potential value of GRACE observations for informing the U.S. Drought Monitor. The GRACE TWS anomaly signal (Figure 9a) provides a good approximation to the large scale extent and severity of the drought conditions recorded by the U.S. Drought Monitor (Figure 9b) in the southeastern United States in August 2007. The GRACE-based DI percentiles of surface soil moisture, root zone soil moisture and groundwater storage (Figure 9c) match the drought monitor map in most places, and contribute additional information on the spatial heterogeneity

Table 3. Evaluation of Ground Water Storage Simulations From Open-Loop (OL2) and Data Assimilation (DAS) Runs Against Measured Groundwater[a]

	Basin 1	2	14	15	16	17	18	19	20	22	26
					Correlation Coefficient r						
Well-OL2	0.70 ± 0.05	0.92 ± 0.02	0.83 ± 0.03	0.82 ± 0.03	0.91 ± 0.02	0.92 ± 0.02	0.85 ± 0.03	0.86 ± 0.03	0.89 ± 0.02	0.69 ± 0.05	0.85 ± 0.04
Well-DAS (20 mm)	**0.86**	0.91	0.85	**0.86**	**0.97**	*0.87*	**0.91**	**0.94**	0.87	0.67	*0.79*
Well-DAS (10 mm)	**0.85**	0.91	0.85	0.84	**0.97**	*0.86*	**0.92**	**0.94**	0.89	0.64	*0.80*
					RMS						
Well-OL2	1.81	1.96	3.21	2.98	5.18	1.00	2.45	4.94	2.65	3.16	3.26
Well-DAS (20 mm)	1.11	2.22	2.35	2.31	2.17	1.26	3.28	2.34	2.49	3.23	3.98
Well-DAS (10 mm)	1.15	2.23	2.29	2.36	2.16	1.29	3.38	2.30	2.40	3.36	3.98

[a]Pearson's correlation coefficient r and RMS error (mm) are calculated for 11 basins across the United States with respect to daily basin-averaged ground water time series based on measurements from 239 monitoring wells (see section 2.8.2). Approximate 95% confidence intervals are given for the OL2 results. On the basis of these, bold font indicates basins with a significant increase in r relative to OL2 at the 5% significance level, whereas italics indicate a significant decrease in r relative to OL2 at the 5% level. Results are shown for a GRACE observation error of 20 and 10 mm, respectively. See Figure 1 for basin denomination.

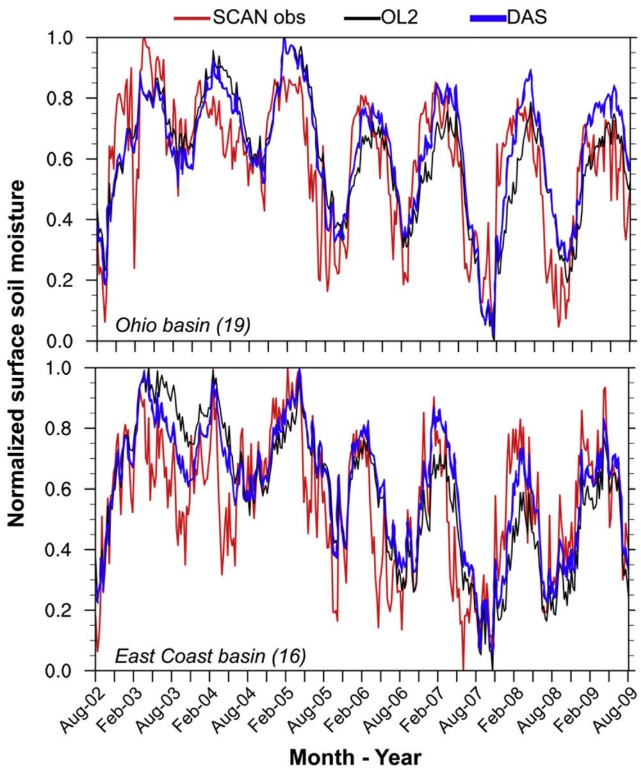

Figure 8. Time series plots of surface soil moisture for the Ohio and East Coast basins comparing weekly SCAN observations with open-loop (OL2) and data assimilation (DAS) simulations.

of droughts. Further, the GRACE-based DIs may help to distinguish hydrological droughts (those that affect deep water storage and streamflow) from agricultural droughts (which affect surface soil moisture and vegetation). Unlike GRACE TWS anomalies (Figure 9a), the DI and drought monitor percentile data are directly comparable, both being representative of drought severity relative to location and season specific dry events.

[43] Figure 10 displays monthly time series drought percentile data for California (a) and for four states (AL, GA, MS, TN) heavily influenced by the 2007 Southeastern drought (b). The plots compare U.S. Drought Monitor output with groundwater storage (gws) drought indicators (DI) on the basis of open-loop simulations (OL2) and model-assimilated GRACE TWS observations (DAS). In CA the groundwater DI generally provides a reasonable approximation to the time series of drought severity levels reported by the U.S. Drought Monitor (Figure 10a). The data assimilation contributes additional skill as evidenced by a significant increase in time series correlation coefficients (r) from 0.71 (OL2) to 0.86 (DAS) (Figure 10a). The RMS difference between the groundwater DI and drought monitor percentiles is also markedly reduced as a result of assimilating the GRACE data (Figure 10a). Improvements are most pronounced toward the end of the time series record where the moderate to severe drought intensities (D1–D2) are better approximated by the GRACE-based DIs. Over this time period, GRACE recorded the largest TWS dry anomaly signals in the fall of 2007 and 2008 (Figure 5). While the

Table 4. Pearson's Correlation Coefficient r of Soil Moisture Simulations From Open-Loop (OL2) and Data Assimilation (DAS) Runs Against Daily Time Series of Soil Moisture Obtained From Soil Climate Analysis Network (SCAN) Sites for Eight Basins Across the United States With Reasonable SCAN Coverage[a]

	Basin							
	1	2	15	16	17	18	19	22
$SCAN_{sfsm}$-$OL2_{sfsm}$	0.38 ± 0.09	0.82 ± 0.03	0.70 ± 0.05	0.73 ± 0.05	0.71 ± 0.05	0.92 ± 0.02	0.76 ± 0.04	0.80 ± 0.03
$SCAN_{sfsm}$-DAS_{sfsm}	0.39	0.82	**0.77**	**0.82**	0.72	*0.89*	**0.81**	0.82
$SCAN_{rtzsm}$-$OL2_{rtzsm}$	0.17 ± 0.12	0.83 ± 0.03	0.72 ± 0.05	0.66 ± 0.06	0.71 ± 0.05	0.87 ± 0.03	0.78 ± 0.04	0.76 ± 0.04
$SCAN_{rtzsm}$-DAS_{rtzsm}	0.20	0.83	**0.77**	**0.77**	0.71	*0.83*	**0.84**	0.77

[a]The 95% confidence intervals are given for the OL2 results. On the basis of these, bold font indicates basins with a significant increase in r relative to OL2 at the 5% significance level, whereas italics indicate a significant decrease in r relative to OL2 at the 5% level. Results are shown for a GRACE observation error of 20 mm. The subscripts sfsm and rtzsm denote surface and root zone soil moisture, respectively. See Figure 1 for basin denomination.

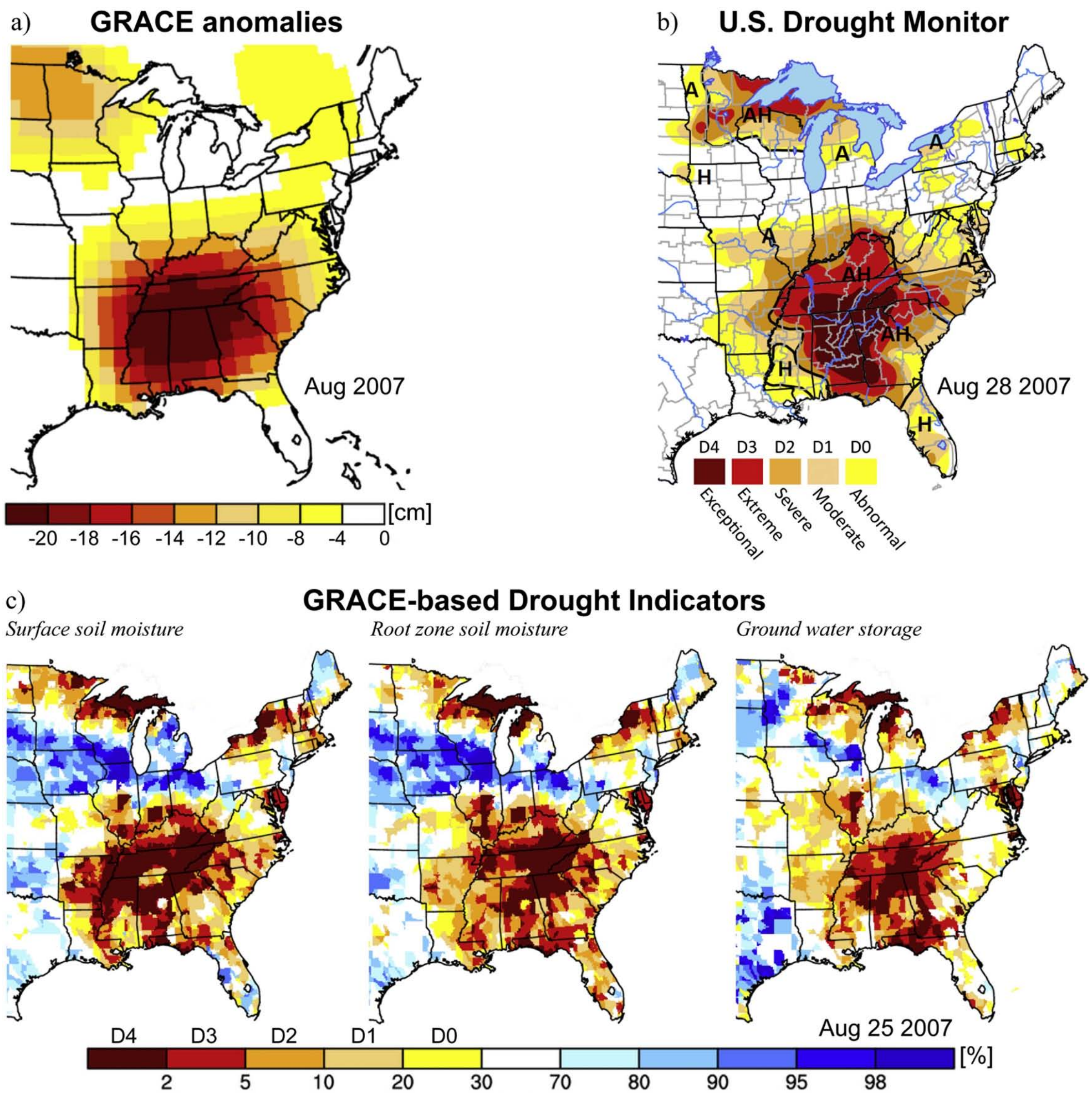

Figure 9. Correspondence between (a) the GRACE monthly water storage anomaly fields, (b) the U.S. Drought Monitor product, and (c) drought indicators based on model-assimilated GRACE TWS observations during the drought in the southeastern United States in August 2007. In Figure 9b A, H, and AH define agricultural drought, hydrological drought, and a mix of A and H, respectively.

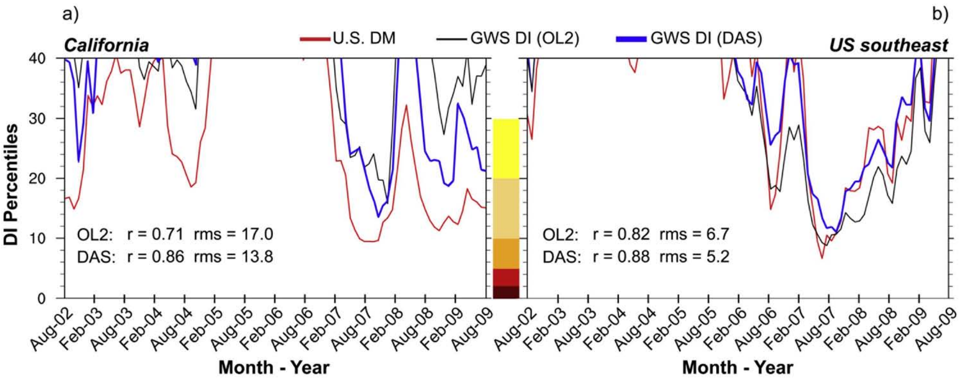

Figure 10. Time series plots of groundwater storage (GWS) drought indicator (DI) and U.S. Drought Monitor percentile data for (a) California and (b) the U.S. Southeast (including the states of Alabama, Georgia, Mississippi, and Tennessee). The GWS DIs were based on open-loop model estimates (OL2) and model-assimilated GRACE TWS observations (DAS). Time series correlation coefficients (r) and root-mean-square errors (RMS) between the model-estimated DIs and the drought monitor output are also shown.

bedrock depth adjustment clearly improved the model's ability to represent drought severity from 2007 onward, the model was still not fully capable of accommodating the GRACE anomalies. Reducing the GRACE observation error only marginally increased the increments during this period. This suggests that increasing the direct perturbation of the prognostic water storage variables may be necessary (section 3.1). The improvement in data assimilation skill observed for CA (Figure 10a) conflicts with the time series correlation results obtained using groundwater measurements as the assumed truth (Figure 7b). As previously noted, the limited spatial coverage and representativeness of the wells in CA may be a reason for the conflicting evidence as the concentration of wells in the southwestern region of the state may not properly reflect the dryness condition of the basin as a whole.

[44] For the four U.S. southeastern states (AL, GA, MS, TN) the statewide averaged groundwater DI percentiles capture the increasing frequency and severity of the drought from 2005 onward (Figure 10b). The agreement between the GRACE-based DI and the drought monitor percentiles is described by a time series correlation coefficient of 0.88, which is a significant improvement over the open-loop results ($r = 0.82$) (Figure 10b). These results are consistent with the evaluation based on groundwater well observations over the eastern United States, which yielded significant improvements in skill because of the assimilation of GRACE data (Figure 7b).

[45] The surface soil moisture and root zone soil moisture DIs were also highly correlated with the drought monitor output for the four U.S. southeastern states with time series correlation coefficients (r_{DAS}) of 0.89 and 0.88, respectively (not shown). However, the impact of the GRACE data assimilation was less pronounced (r_{OL2} of 0.87 and 0.86, respectively). In CA, the ability of the soil moisture DIs to capture drought conditions reported by the U.S. Drought Monitor was significantly reduced in comparison with the groundwater DI with characteristic time series r on the order of 0.56 (not shown). This suggests a closer coupling between the surface and groundwater compartments in the eastern United States and a generally greater drought detection value of the groundwater DI, at least with respect to the information used to draw the drought monitor maps. This is not surprising, as the effects of drought may prevail long after climatic conditions have improved because of the slow process of aquifer recharge, while soil moisture responds rapidly to short-term variations in atmospheric forcing.

4. Discussion and Conclusions

[46] The GRACE data assimilation system (GRACE DAS) was applied to North America as part of a larger effort to demonstrate that drought conditions can be identified more accurately and objectively by integrating spatially, temporally, and vertically disaggregated GRACE data into the U.S. and North American Drought Monitor products, substituting for ground-based observations of groundwater and soil moisture which are currently lacking. The specific objectives of this paper were to describe the methodology behind the development of drought indicators (DI) based on model-assimilated GRACE terrestrial water storage (TWS) data and to assess improvements in hydrological modeling skill and drought detection resulting from the assimilation of GRACE TWS data.

[47] The monthly production frequency and coarse spatial resolution of GRACE TWS fields limit their utility for applications that require near-real-time input with fine spatial and temporal resolutions. This study confirms that data assimilation may be the key to realizing the full potential of GRACE TWS anomalies for hydrological applications, as it facilitates spatial and temporal downscaling, extrapolation to near real time, and vertical stratification into groundwater and soil moisture components, which individually are more useful for scientific applications.

[48] By increasing the water storage capacity of the Catchment land surface model we made it more accommodating of GRACE TWS anomalies during times of severe drought. This was accomplished by increasing the depth to bedrock, and it generally improved the model's ability to represent the magnitude and seasonality of observed TWS and groundwater storage. Using an extensive data set of groundwater observations from monitoring wells across the U.S. as "truth", the skill of modeled groundwater storage

was shown to be significantly improved in major parts of the United States as a result of assimilating the GRACE observations. Negative skill scores (significant decrease in skill associated with assimilating GRACE) in the Missouri and California basins are likely attributable to spatial undersampling, with existing monitoring sites being unable to properly capture the prevailing drought conditions representative of the basins as a whole. Despite a limited number of in situ soil moisture observation sites, positive skill scores (significant increase in skill associated with assimilating GRACE) particularly in the Eastern U.S. provided evidence that the GRACE TWS data contain useful independent information that the assimilation algorithm was able to translate into superior soil moisture estimates.

[49] A 62 year bias-corrected climatology of model simulated soil moisture and groundwater was used to convert GRACE-assimilated moisture fields into drought indicator (DI) percentiles for potential integration into the U.S. and North American Drought Monitors. Initial comparisons between the model-based drought indicators and U.S. Drought Monitor output indicated improved correlations attributable to the assimilation of GRACE TWS data. Thus GRACE data assimilation helps to overcome limitations of land surface models, which include imperfect soil, vegetation, and topographical parameters, errors in the meteorological forcing data, and simplified or incomplete representation of surface water and energy processes such as lateral surface and groundwater flows. The GRACE-based groundwater storage drought indicator proved to be particularly useful, reflecting the longer-term meteorological anomalies associated with drought. The results highlight the potential value of drought indicators based on model-assimilated GRACE TWS observations for identifying drought conditions more comprehensively and objectively.

[50] One source of uncertainty is the contribution of groundwater mining (withdrawals in excess of net recharge) to groundwater storage variations. Water management and groundwater withdrawals are not simulated by CLSM (which may in fact be advantageous for assessing natural impacts of drought), but their effects are detected by GRACE. Thus GRACE DAS likely indicates some level of groundwater depletion in overappropriated aquifers that is not caused (directly) by drought. For example, the High Plains aquifer of the central United States and the Central Valley aquifer in California are known to have experienced water level declines over the last several decades because of groundwater abstraction for irrigation. We believe the adverse consequences of groundwater abstraction impacts for the GRACE-based drought indicators have so far been minor, because the drought indicators in the areas of the overappropriated aquifers have yet to show wetness percentiles that are significantly lower than those of the U.S. Drought Monitor. An explanation for this is that the mean rate of anthropogenic groundwater storage decline at the river basin scale of the GRACE observations as applied here is at least an order of magnitude smaller than the natural seasonal to interannual variability of groundwater storage. Nevertheless, it is an issue that must continue to be monitored.

[51] More in depth analyses are ongoing to establish the benefit of the GRACE data assimilation for drought monitoring activities across the nation and the North American domain as a whole. As part of this process near-real-time GRACE-based drought indicators are currently being made available to the U.S. and North American Drought Monitor community for independent assessment of their value for informing operational drought monitor products (http://www.drought.unl.edu/MonitoringTools/NASAGRACEDataAssimilation.aspx). Furthermore, the added benefit of incorporating the GRACE-based drought indicators will be analyzed through comparison with the current suite of short- and long-term objective indicators and by correlating detectable differences at the regional and local scale to the final U.S. and North American Drought Monitor products.

[52] **Acknowledgments.** This project was supported by NASA's Applied Sciences program. VIC data were processed and kindly made available by Huilin Gao. GRACE land data were processed by Sean Swenson, supported by the NASA MEASURES Program, and are available at http://grace.jpl.nasa.gov.

References

Boronina, A., and G. Ramillien (2008), Application of AVHRR imagery and GRACE measurements for calculation of actual evapotranspiration over the Quaternary aquifer (Lake Chad basin) and validation of groundwater models, *J. Hydrol.*, *348*, 98–109.

Chambers, D. P., J. Wahr, M. E. Tamisiea, and R. S. Nerem (2010), Ocean mass from GRACE and glacial isostatic adjustment, *J. Geophys. Res.*, *115*, B11415, doi:10.1029/2010JB007530.

Chen, J. L., C. R. Wilson, D. Blankenship, and B. D. Tapley (2009a), Accelerated Antarctic ice loss from satellite gravity measurements, *Nat. Geosci.*, *2*, 859–862, doi:10.1038/NGEO694.

Chen, J. L., C. R. Wilson, B. D. Tapley, Z. L. Yang, and G. Y. Niu (2009b), 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models, *J. Geophys. Res.*, *114*, B05404, doi:10.1029/2008JB006056.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a GCM: 2. Parameter estimation and model demonstration, *J. Geophys. Res.*, *105*, 24,823–24,838.

Forman, B. A., R. H. Reichle, and M. Rodell (2011), Assimilation of terrestrial water storage from GRACE in a snow-dominated basin, *Water Resour. Res.*, *48*, W01507, doi:10.1029/2011WR011239.

Gao, H., Q. Tang, C. R. Ferguson, E. F. Wood, and D. P. Lettenmaier (2010), Estimating the water budget of major US river basins via remote sensing, *Int. J. Remote Sens.*, *31*(14), 3955–3978, doi:10.1080/01431161.2010.483488.

Güntner, A. (2008), Improvement of global hydrological models using GRACE data, *Surv. Geophys.*, *29*, 375–397, doi:10.1007/s10712-008-9038-y.

Hasegawa, T., Y. Fukuda, and K. Yamamoto (2009), The 2006 Australian drought detected by GRACE, in *From Headwaters to the Ocean*, edited by M. Taniguchi et al., pp. 363–367, Taylor and Francis, London.

Klees, R., E. A. Zapreeva, H. C. Winsemius, and H. H. G. Savenije (2006), The bias in GRACE estimates of continental water storage variations, *Hydrol. Earth Syst. Sci. Discuss.*, *3*, 3557–3594.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord (2008), Implementation of a new 3DVAR analysis as part of the NCEP global data assimilation system, *Weather Forecast.*, *24*, 1691–1705.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *J. Geophys. Res.*, *105*, 24,809–24,822.

Lawrimore, J., R. Heim, M. Svoboda, V. Swail, and P. J. Englehart (2002), Beginning a new era of drought monitoring across North America, *Bull. Am. Meteorol. Soc.*, *83*(8), 1191–1192.

Leblanc, M. J., P. Tregoning, G. Ramillien, S. O. Tweed, and A. Fakes (2009), Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, *Water Resour. Res.*, *45*, W04408, doi:10.1029/2008WR007333.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, *J. Geophys. Res.*, *99*, 14,415–14,428.

Liu, Q., et al. (2011), The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system, *J. Hydrometeorol.*, *12*, 750–765, doi:10.1175/JHM-D-10-05000.1.

Lo, M.-H., J. S. Famiglietti, P. J.-F. Yeh, and T. H. Syed (2010), Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data, *Water Resour. Res.*, *46*, W05517, doi:10.1029/2009WR007855.

Luthcke, S. B., A. A. Arendt, D. D. Rowlands, J. J. McCarthy, and C. F. Larsen (2008), Recent glacier mass changes in the Gulf of Alaska region from GRACE mascon solutions, *J. Glaciol.*, *54*(188), 767–777.

Mishra, A. K., and V. P. Singh (2010), A review of drought concepts, *J. Hydrol.*, *391*, 202–216, doi:10.1016/j.jhydrol.2010.07.012.

Niu, G.-Y., K.-W. Seo, Z.-L. Yang, C. Wilson, H. Su, J. Chen, and M. Rodell (2007), Retrieving snow mass from GRACE terrestrial water storage change with a land surface model, *Geophys. Res. Lett.*, *34*, L15704, doi:10.1029/2007GL030413.

Ramillien, G., F. Frappart, A. Güntner, T. Ngo-Duc, A. Cazenave, and K. Laval (2006), Time variations of the regional evapotranspiration rate from Gravity Recovery and Climate Experiment (GRACE) satellite gravimetry, *Water Resour. Res.*, *42*, W10403, doi:10.1029/2005WR004331.

Reichle, R. H., and R. D. Koster (2003), Assessing the impact of horizontal error correlations in background fields on soil moisture estimation, *J. Hydrometeorol.*, *4*, 1229–1242.

Reichle, R. H., and R. D. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.*, *31*, L19501, doi:10.1029/2004GL020938.

Rodell, M., and J. S. Famiglietti (2001), An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *37*, 1327–1340.

Rodell, M., J. S. Famiglietti, J. Chen, S. I. Seneviratne, P. Viterbo, S. Holl, and C. R. Wilson (2004), Basin scale estimates of evapotranspiration using GRACE and other observations, *Geophys. Res. Lett.*, *31*, L20504, doi:10.1029/2004GL020873.

Rodell, M., B. F. Chao, A. Y. Au, J. Kimball, and K. McDonald (2005a), Global biomass variation and its geodynamic effects, 1982–1998, *Earth Interact.*, *9*, 1–19.

Rodell, M., P. R. Houser, A. A. Berg, and J. S. Famiglietti (2005b), Evaluation of 10 methods for initializing a land surface model, *J. Hydrometeorol.*, *6*, 146–155.

Rodell, M., J. Chen, H. Kato, J. S. Famiglietti, J. Nigro, and C. R. Wilson (2007), Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE, *Hydrogeol. J.*, *15*(1), 159–166, doi:10.1007/s10040-006-0103-7.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*, 999–1003, doi:10.1038/nature08238.

Rowlands, D. D., S. B. Luthcke, S. M. Klosko, F. G. Lemoine, D. S. Chinn, J. J. McCarthy, C. M. Cox, and O. B. Andersen (2005), Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, *Geophys. Res. Lett.*, *32*, L04310, doi:10.1029/2004GL021908.

Schmidt, R., et al. (2006), GRACE observations of changes in continental water storage, *Global Planet. Change*, *50*(1–2), 112–126, doi:10.1016/j.gloplacha.2004.11.018.

Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, *J. Clim.*, *19*(13), 3088–3111.

Sheffield, J., C. R. Ferguson, T. J. Troy, E. F. Wood, and M. F. McCabe (2009), Closing the terrestrial water budget from satellite remote sensing, *Geophys. Res. Lett.*, *36*, L07403, doi:10.1029/2009GL037338.

Stieglitz, M., A. Ducharne, R. Koster, and M. Suarez (2001), The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales, *J. Hydrometeorol.*, *2*(3), 228–242.

Strassberg, G., B. R. Scanlon, and D. Chambers (2009), Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United Sates, *Water Resour. Res.*, *45*, W05410, doi:10.1029/2008WR006892.

Su, H., Z. L. Yang, R. E. Dickinson, C. R. Wilson, and G.-Y. Niu (2010), Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information, *J. Geophys. Res.*, *115*, D10104, doi:10.1029/2009JD013035.

Svoboda, M., et al. (2002), The drought monitor, *Bull. Am. Meteorol. Soc.*, *83*(8), 1181–1190.

Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.*, *33*, L08402, doi:10.1029/2005GL025285.

Swenson, S., and J. Wahr (2009), Monitoring the water balance of Lake Victoria, East Africa, from space, *J. Hydrol.*, *370*, 163–176, doi:10.1016/j.jhydrol.2009.03.008.

Swenson, S., P. J.-F. Yeh, J. Wahr, and J. S. Famiglietti (2006), A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, *Geophys. Res. Lett.*, *33*, L16401, doi:10.1029/2006GL026962.

Swenson, S., J. Famiglietti, J. Basara, and J. Wahr (2008), Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, *Water Resour. Res.*, *44*, W01413, doi:10.1029/2007WR006057.

Syed, T. H., J. S. Famiglietti, J. Chen, M. Rodell, S. I. Seneviratne, P. Viterbo, and C. R. Wilson (2005), Total basin discharge for the Amazon and Mississippi River basins from GRACE and a land-atmosphere water balance, *Geophys. Res. Lett.*, *32*, L24404, doi:10.1029/2005GL024851.

Syed, T. H., J. S. Famiglietti, and D. Chambers (2008), GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales, *J. Micrometeorol.*, *10*, 22–40, doi:10.1175/2008JHM993.1.

Tang, Q., H. Gao, P. Yeh, T. Oki, F. Su, and D. P. Lettenmaier (2010), Dynamics of terrestrial water storage change from satellite and surface observations and modeling, *J. Hydrometeorol.*, *11*, 156–170, doi:10.1175/2009JHM1152.1.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004), GRACE measurements of mass variability in the Earth system, *Science*, *305*, 503–505.

Velicogna, I. (2009), Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE, *Geophys. Res. Lett.*, *36*, L19503, doi:10.1029/2009GL040222.

Verdin, K., and J. Verdin (1999), A topological system for the delineation and codification of the Earth's river basins, *J. Hydrol.*, *218*, 1–12.

Wahr, J., M. Molenaar, and F. Bryan (1998), Time-variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, *103*, 30,205–30,230.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004), Time-variable gravity from GRACE: First results, *Geophys. Res. Lett.*, *31*, L11501, doi:10.1029/2004GL019779.

Wahr, J., S. Swenson, and I. Velicogna (2006), Accuracy of GRACE mass estimates, *Geophys. Res. Lett.*, *33*, L06401, doi:10.1029/2005GL025305.

Webb, R., C. Rosenzweig, and E. Levine (1991), A global data set of soil particle size properties, *NASA Tech. Memo.*, *4286*, 1–29.

Werth, S., A. Guntner, S. Petrovic, and R. Schmidt (2009), Integration of GRACE mass variations into a global hydrological model, *Earth Planet. Sci. Lett.*, *277*(1–2), 166–173, doi:10.1016/j.epsl.2008.10.021.

Western Governors' Association (2004), Creating a drought early warning system for the 21st century: The National Integrated Drought Information System, report, 13 pp., Denver, Colo. [Available at http://www.westgov.org/.]

Wilhite, D. A. (2000), Drought planning and risk assessment: Status and future directions, *Ann. Arid Zone*, *39*(3), 211–230.

Xie, P. P., and P. A. Arkin (1997), Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bull. Am. Meteorol. Soc.*, *78*, 2539–2558.

Yeh, P. J.-F., S. C. Swenson, J. S. Famiglietti, and M. Rodell (2006), Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *42*, W12203, doi:10.1029/2006WR005374.

Yirdaw, S. Z., K. R. Snelgrove, and C. O. Agboma (2008), GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian Prairie, *J. Hydrol.*, *356*, 84–92.

Zaitchik, B. F., M. Rodell, and R. H. Reichle (2008), Assimilation of GRACE terrestrial water storage into a land surface model: Results for the Mississippi River basin, *J. Hydrometeorol.*, *9*, 535–548.